ABSTRACT

A Collection Processing System enables people to perform complex computational operations on large sets of collections, using a simple and convenient syntax for collection symbolic job requests. A collection symbolic job request is comprised of a user-defined symbolic task name, and a collection reference expression that can refer to large sets of collections. Collections are data-typed sets of computer files that can be manipulated as a set, rather than as individual files. In operation, a Collection Processing System receives a collection symbolic job request from a request originator, and expands the symbolic job request into a list of executable commands that carry out the computational intent of the symbolic job request. First-level symbolic task names are expanded into sequences of second-level task part statements, and then into third-level executable computer commands that are dynamically generated into customized, optimal makefiles. Collection reference expressions are expanded into lists of job triplets comprised of a particular collection name, a computing platform name, and a processing dependency visit order value. The present invention applies third-level executable commands to job triplets to carry out the original collection symbolic job request, in a distributed, multiplatform, scalable way that was not previously known to the art.